Chapter 1
INTRODUCTION

In this thesis, I rely heavily on the idea of epistasis, a difficult term with substantial historical weight. In some respects the history of epistasis and the history of genetics are the same. In this Introduction, I give what I consider to be the relevant historical background for understanding my use of the term, present various definitions for it, and introduce the reader to a body of research concerning the effects of epistasis on what we know about biology. Finally, I connect these ideas to my own research, in which I argue that selectively injecting epistatic considerations into experimental and theoretical models–specifically, using intuitions derived from known biological interactions–can both dramatically increase our understanding of the heritability of traits and refine our ability to understand and predict evolutionary patterns.

1.1 Historical notes on heredity, genotypes, and phenotypes

The first decades of the 20th century were an exciting time for genetics. Mendel's work had been rediscovered, Galton's was never forgotten, and the debate between Darwinists and Lamarckists was waged with increasingly precise experimental tests. The crucial question under study was the mechanism of heredity; that is, how are observable differences in character among organisms propagated across generations [115]? For example, Galton chose to investigate the relative roles of 'nature' and 'nurture' by studying the characteristics of twins [73], and Johannsen self-fertilized crop plants to obtain genetically stable 'pure lines' where the same qualities could be studied more exactly [116]. These investigations almost unconsciously led to a further question: how does the ontogeny of an organism give rise to a character? Specifically, immediately following Jo-

hannsen, Woltereck used 'pure lines' of Daphnia to show that specific morphological changes could be reproducibly achieved by either manipulating growth conditions or substituting different isogenic Daphnia lines [292].

The first question, of heredity, is by far the easier, having been answered in formal terms by the succeeding century of research into chromosome theory, genetic mapping, and a litany of other inheritance mechanisms. The second question, being much more open-ended, must be answered anew in each case. For instance, Woltereck's observation that the head-height of *Hyalodaphnia cucullata* may be increased by a combination of heat and rich food is unlikely to generalize to the articulation of finger bones in humans, though in both cases heritable variation in these traits may be attributed to definite loci on inherited chromosomes in the respective organisms. However, I would argue that a meaningful answer to the first question (heredity) is not very helpful without some insight into the second (physiological mechanism).

However, for many years our ability to ascertain such mechanism was essentially nil with respect to the molecular activity of the heritable material itself. It was generally agreed that some chemical activity probably distinguished genes [115], but the majority of important work was determined by studying segregation ratios from crosses, rather than the physiological basis of phenotypes. Nonetheless, the ideas of 'genes' and of 'genotypes' provided rich material for early geneticists, assisting the resolution of quite complicated segregation patterns. Specifically, understanding the gene as a discrete locus with influence upon a character (or 'phenotype') allowed the development of Mendelian thought. The biometricians (such as Galton) were, in contrast, interested in exploring the phenomena underlying continuous variation in characters. This view of hereditary variation was apparently in conflict with the Mendelian model of a few discrete packets of genetic information. Out of this debate, from the Mendelian side, came the now-familiar idea of the 'gene' (a discrete genetic determinant of a character), the 'genotype' (a fixed complement of genes in a given organism), and the 'phenotype' (the directly observable character which can be measured upon a given organism) [115].

Fisher put an end to this dispute with a mathematical framework [68] showing that Mendelian segregation of genes could in principle lead to the continuous variation in phenotypes observed by the biometricians.

1.1.1 Fisher's innovations.

For the purposes of this thesis, I will emphasize some relevant conceptual changes ushered in by Fisher's quantitative genetic framework in his 1918 paper [68] and subsequent work.

Fisher on quantitative genetics

First, Fisher implicitly assumed there is some direct, biologically meaningful mapping between variation in phenotypes and genotypes, such that variation in the phenotype is decomposable into quantities attributable to specific genes. This anticipated the 'genotype-phenotype map' concept popularized later [3]. Interestingly, this direct abstraction of the genotype idea had previously been explicitly discouraged by Johannsen [115], who cautioned that such a leap was dangerous in ignorance of the actual hereditary material and the mechanisms by which phenotypes were generated from the hereditary material. However, at the time the resolution of the Mendelian/biometrician divide was too desirable to be laid aside for such misgivings. In consequence, Fisher's framework dealt with idealized, purely abstract genes, whose existence and influence had more to do with mathematical convenience than with direct observation.

Second, Fisher used a series of assumptions about the structure of populations, the number of relevant genes, and the way that genes work together to create a mathematically tractable model of how phenotypes are created [182]. Specifically, he assumed that the number of genes contributing to any phenotype was large, with relatively small contributions from each gene. When this is the case, and the population of organisms tends to infinity in size, then the phenotype in question will be normally distributed across the population.

Together with the concept of the genotype mentioned above, Fisher shows that the normally distributed phenotypic variance (written σ_P^2) can be decomposed into independent portions attributable to each Mendelian 'factor' or gene i (among n genes total) and to a non-genetic error term (e):

$$\sigma_P^2 = \sum_{i=1}^{n} \sigma_i^2 + \sigma_e^2 \tag{1.1}$$

This independence between genes leads to the property called 'additivity', because the genetic variance of the phenotype can be computed as a simple linear combination $\sigma_a^2 = \sum_{i=1}^{n} \sigma_i^2$ of the variance attributable to each gene. Similarly, the expected value of the phenotype can be computed as a linear combination of the effects of each gene[1]:

$$E[Phenotype] = \alpha + \sum_{i=1}^{n} \beta_i G_i + \epsilon \tag{1.2}$$

Where there are n genes, G_i is the a 0/1 indicator of an alternate allele at the i-th locus (or gene), β_i is the the effect attributable to the alternate allele at the i-th locus, α is the intercept term (corresponding to the phenotype when all loci G_i take the value 0), and ϵ is the error introduced by all other factors (sampling error, measurement error, environmental variation). For simplicity, this example considers only a haploid system. With many independent genes, this decomposability yields a series of predicted phenotypic correlations between relatives of different degree within the population. These phenotypic correlations are related to the 'heritability' of phenotypes (sometimes written h^2), or the proportion of phenotypic variation that is attributable to genetic variation ($h^2 = \frac{\sigma_a^2}{\sigma_P^2}$). Estimates of these correlations were the real object of Fisher's study, for formulating a Mendelian mathematical basis for the well-known phenotypic resemblance between relatives.

[1]The framework laid out in this 1918 paper is alternately famous for introducing one of the most popular statistical methods, the analysis of variance (ANOVA). The idea of decomposing variances was generalizable to any problem in detecting associations between a quantitative normally-distributed dependent variable and discrete factorial variables. ANOVA has subsequently seen wide usage beyond genetics. Certain properties of ANOVA, specifically its deficiencies in jointly estimating main effects and interaction terms, have been criticized both in applied statistics [302] and quantitative genetics [146].

Notably, Fisher also treated classes of phenotype-controlling genetic variation that would not show up in correlations among relatives due to distortions. These were interactions among genes, or 'epistasis', which violated Fisher's assumption about the independence of genes. To introduce epistasis into the model of Equation 1.2, we can add terms corresponding to the interactions between each pair of genes:

$$E[Phenotype] = \alpha + \sum_{i=1}^{n} \beta_i G_i + \sum_{i=1}^{n} \sum_{j \neq i}^{n} \beta_{i,j} G_{i,j} + \epsilon. \quad (1.3)$$

Where $G_{i,j}$ and $\beta_{i,j}$ correspond, respectively, to an indicator for the joint genotype at loci i and j, and the effect attributable to the interaction (which can be non-symmetric). Fisher's variance decomposition could in turn be modified to yield a term corresponding to the genetic variance from epistasis [38]. I shall discuss epistasis in more detail below, and simply note here that Fisher was not particularly interested in this variance component other than as a nuisance parameter similar to ϵ; an acceptable loss similar to error in measuring the phenotype [182].

Thus, Fisher emphasizes the estimation of phenotype based on only the readily tractable component of genotypic variation, and does not claim to provide a causal model by which specific genes (as opposed to ideal Mendelian factors) influence phenotypes. In this, Fisher followed the example of the preceding Mendelians, for whom the purely hypothetical genotype was much less interesting than the ability to dissect phenotypic variation in terms of specific discrete factors varying between monolithic genetic varieties [115]. However, as time went on, geneticists tended to replace the mathematical abstractions of genes in Fisher's model with alleles actually isolated in nature.

Fisher on evolutionary genetics

In later work, Fisher exploited his ideas on quantitative genetics to derive a mathematical theory of how these hereditary principles would behave in an evolutionary setting [69]. Throughout, he relied strongly on Darwin's intuition that, in order for natural

selection to operate on phenotype, phenotypic variation must be heritable [290]. Otherwise, selection will be ineffective at promoting phenotypic change from generation to generation. Fisher's prior work provided tools by which heritability could be directly estimated. In consequence, Fisher came to equate the rate of evolutionary change in a population with the heritability of the phenotype under selection. Stated differently, selection uses up genetic variation to effect phenotypic change. For instance, the change of a phenotype Z in response to a selection on individuals with specific values of Z can be written

$$\Delta Z = h^2 S, \tag{1.4}$$

where h^2 is the heritability of Z, and S signifies the intensity of selection [67]. In the ideal case, the trait of 'fitness', or reproductive success, is substituted for the trait subject to quantitative analysis in Fisher's earlier work.

This system provides an intuitive formulation by which evolution by selection can proceed within a population according to Mendelian principles of segregating factors controlling phenotypic variation, though once again the actual identities of these factors were in practice irrelevant. For example, if a population of organisms with a trait Z with average value $\overline{Z}$ is subject to a selection under which the selected subpopulation Z' has mean $\overline{Z'}$, then we can rewrite equation 1.4 to give the expected phenotype of the next generation Y:

$$E[Y] = \overline{Z} + h^2(\overline{Z'} - \overline{Z}) \tag{1.5}$$

Thus showing a direct relationship between phenotypic change across generations and the genetic variance, and thereby showing how Mendelian factors, summarized in h^2, can contribute to evolutionary change (Fisher called his formulation of this direct genetic response to selection the "Fundamental Theorem of Natural Selection"). Of course, under the standard model, this selection will reduce h^2 in the next generation by changing the allele frequencies of the population, which will now be somewhat biased towards the alleles responsible for the difference ΔZ by a quantity proportional

to the initial h^2 [293, 294]. It is in this sense that genetic variance can be used up in selection.

The assumptions of this framework are the same strong assumptions made for Fisher's quantitative genetics framework, concerning very large panmictic populations, where fitness is determined by a large number of independently-contributing genes. For instance, the h^2 mentioned above is sometimes called the 'narrow-sense' heritability, in that it includes only additive genetic variation (σ_a^2), as opposed to the more inclusive 'broad-sense heritability', which explicitly includes non-additive genetic phenomena such as epistasis and dominance.

As previously mentioned, Fisher assumes that additivity is a satisfactory model for describing the relationship of genotypes, phenotypes, and fitness. In a Fisherian world, fitness can be visualized as a smooth hill that one is climbing, where the peak of the hill is the fitness optimum, and higher elevation indicates higher fitness. Because each step up the hill increases fitness, one is guaranteed to reach the peak if one always chooses paths of higher fitness; there is only one peak, and all paths lead monotonically from lower fitness to optimal fitness.

1.1.2 Wright's problems.

The analogy of fitness as a landscape in which one climbs is generally traced back to Fisher's contemporary Sewall Wright. Wright was dissatisfied with the idea of a single stable peak of fitness for a population, drawing on both theory of his own devising and multiple empirical examples from the literature [293]. While at its root the argument came back to the problem of epistasis in quantitative genetics, the immediate and more famous disputes between Fisher and Wright centered on models for evolutionary genetics, and specifically population structure.

Fisher once wrote to Wright (quoted in [283]), "...I believe that N [population size] must usually be the total population on the planet [of the organism in question]..." This is obviously an expansive view of the idea of a population. For example, it is unclear to

what degree distinguishable local varieties of species should be included.

Thus, Fisher dealt with many mathematical complexities of evolutionary theory by simply assuming very large N ($\frac{1}{4N} \ll s$ where s is the selection coefficient, and $Nm \gg 1$ where m is the rate of migration among populations, $Nu \gg 1$ where u is the mutation rate) [69]. Wright investigated the consequences of varying N extensively outside of these boundaries [293], finding qualitatively quite different behaviors among models. Specifically, selection acts relatively slowly in both very large populations (because of the extremely long time to fixation of beneficial alleles) and in very small populations even strong selection can fail to favor beneficial alleles (when $s < \frac{1}{N}$). Although Fisher favored the first of these scenarios, Wright judged that neither was fast enough to yield the observed diversity of life outside of special cases such as artificial selection.

Consequently, the more interesting cases are when $s \not\gg u \not\gg \frac{1}{4N}$, an intermediate region of parameter space where selection can still act, but where genetic drift is strong enough to lead to substantial variation in gene frequencies under selection. Furthermore, the time required for a favorable allele to go to fixation will be somewhat smaller in a smaller population.

A crucial difference which must here be noted between Fisher and Wright is the emphasis that Wright placed on gene interactions, or epistasis as discussed above. While Fisher's framework will function well in cases where the causal relationship between genotype and phenotype is more or less additive, in more complicated epistatic situations (which even then were known to be common, due to work on incompatibility between varieties [263, 85, 129]), adaptive evolution by selection according to a Fisherian model will be difficult at best. Wright points out that if genes are selected in combinations rather than individually, the combinatorics of finding favorable genotype combinations can become very hard for a single large population [294]. That is, the additive, average effect of an allele may change with its frequency or the background in which it resides, leading to many false starts in adaptation [290].

For these reasons, Wright considered the Fisherian model of one large population to

be too inefficient to generate observed evolutionary change, even at permissive values of the population genetic parameters. He therefore considered the model that species exist in dispersed, loosely defined populations in a 'metapopulation' network, where individuals generally stay in the same small population, but occasionally emigrate. Wright argued that this approach, of allowing many replicated noisy evolutionary experiments with small amounts of mixing, would more efficiently find favorable gene combinations, thanks to the random assortment of alleles. A pertinent feature of this scheme is that small populations would tend to fix alleles, and fix them differently across populations, allowing other alleles in epistasis with fixed alleles to behave more 'additively' [283].

In support of this model, Wright cites examples from plant breeding, under which phenotypic change is most quickly achieved by self-fertilizing individual segregants (i.e. restricting population size) in alternation with outcrossing [293]. This contradicts Fisher's model, under which one would expect the greatest returns to selection by maximizing population size and thus making selection on additive variation more efficient. More recent explicit tests with insect metapopulations have tended to support Wright's predictions [282, 250].

1.2 The many names of epistasis.

We shall now finally deal with epistasis, the central idea of this thesis. Epistasis is famously ill-defined [207, 208], but it can be considered a rough-and-ready word to capture the influence of interactions between genetic loci. These interactions could take the form of physical associations between encoded gene products or residues, the joint action of two loci on a trait, or the relative ordering of genes in genetic pathways. By all accounts, cytological interactions between gene products occur at a massive scale [133], and gene interactions are crucial for organismal viability under laboratory-derived mutant analysis [47]. From these observations, one might naively expect that epistasis would be unambiguously important. However, the importance of epistasis for phenotypic variation in natural populations is quite controversial, with some claiming it to be

negligible [100] and others claiming it to be all-important [102].

We are left in a curious situation where, despite the overwhelming mechanistic importance of interactions, there are cogent and rigorous arguments for the unimportance of epistasis for heritable variation [225]. This brings us back again to the old duality exposed by Woltereck [292], in distinguishing between the biological mechanisms by which phenotypes are generated and the statistical description of heritable phenotypic variation. Specifically, the argument concerns whether additive models are adequate for understanding the diversity of phenotypes within and across species. Mathematically, epistasis is defined as a quantitative departure from the expectations of additivity [33, 211]. For these reasons, it shall be instructive to briefly examine the idea of additivity before moving on to different definitions of epistasis relevant to this thesis.

1.2.1 The uses of additivity.

As may be seen in Equation 1.2, additive models consider that genes contribute to phenotype independently of one another, such that the effects of two loci are in no way dependent upon one another[2]. It has been widely acknowledged by both Fisher and his latter-day adherents that this model is a convenient approximation rather than truth, and widely defended as such on the basis of parsimony [68, 289, 290, 283, 250, 225]. Certainly, comparing Equation 1.2 to Equation 1.3, one will immediately notice that the pairwise epistatic terms in the model of phenotype increase in the order of the square of n, the number of loci, whereas the additive terms are equal in number to n. The inclusion of interactions of higher order (3-way or 4-way in addition to pairwise) makes matters worse. In any non-trivial example, this can lead to an alarming number of parameters of the model, leading potentially to problems in both estimation and interpretation of the parameters. These are good arguments against exhaustively considering epistatic terms if one can reasonably avoid it.

[2]This is true of normally distributed quantitative phenotypes; in contrast, for log-normally distributed phenotypes, the analogous multiplicative model is appropriate.

Since the assumption of additivity is so widely used and defended, it is worth examining its meaning in the context of quantitative genetics. So let us consider some specific thought experiments, featuring an allele *FBN1** of the *FBN1* gene (disrupting a splice site). Among other pleiotropic phenotypes, this allele acts as an additive contributor to human height on the order of 10cm (a large effect) [183, 225].

1. The allele *FBN1** was identified in human populations, probably of largely European descent given the geography of the research group that discovered several such alleles [183]. Among people of European descent, this allele then changes height by 10cm. If it is truly additive, genetic background is irrelevant, and thus we may expect the same effect in humans from other populations, and probably in Neanderthal or Denisovan individuals. To venture further, this allele would have the same 10cm effect (or a similar scaled transformation) in any metazoan, or plant, or bacterium.

2. The quantitative effects of allele *FBN1** must be independent of alleles at any other locus. There are many known alleles segregating in the human population with potentially refractory phenotypic effects, such as recessive lethal mutations, but if *FBN1** is additive, it necessarily increases adult height by 10cm even in such cases.

These vignettes are obviously specious, and no one would believe in the existence of such effects. Nonetheless, they may illustrate what we mean when we discuss additivity. The first thought experiment may be dismissed on a technicality: any estimate of an additive main effect is a local estimate specific to a population. Thus, the 10cm estimate is specific to the population of humans of European descent in which it was measured. This additive estimate should thus not be applied to any other group of humans or any other species without first re-measuring it in the group in question[3].

[3]Note that this argument implies that estimates should not be extended to any humans other than those actually sampled

The second case is somewhat trickier to dismiss. Unlike the first, there is no objection regarding the species concept, and the potentially epistatic alleles are all actually in the same population in the real world. There may still be a case against it, but it is hard to argue that the hypothetical *FBN1** effect would behave additively in the strict sense. It may follow then that there is no allele of any meaningful effect that is completely independent of the genetic background (alleles with no effect are necessarily independent). Obviously, no one has ever argued otherwise. Instead, as mentioned above, appeals are generally made to the parsimony and adequacy of additive models, rather than their accurate portrayal of causality (as is often repeated, "all models are wrong, some models are useful").

The most sophisticated argument in favor of additive models, which was made by Fisher himself, is that epistatic (and other non-additive) causal effects are modeled as part of the additive variance [68, 33, 284, 208, 225, 172]. In any given population, epistatic effects of a locus will have some average effect across the genotypes in the population, which can be approximated as part of the additive effect of that locus. With large sample sizes and a representative sample of genotypes and environmental conditions, this will lead to quite accurate phenotypic predictions, as stressed by the quote from Fisher above and by others [100]. There are of course many examples of loci contributing to traits via epistasis that show negligible additive effects [25, 296, 208, 138], but these appear to be generally ignored. For this and other reasons, some have presented theoretical arguments that the misspecification of epistatic variance into additive components is an unacceptable tradeoff [284, 182]. However, the defenders of the additive paradigm counter that the parsimony and explanatory adequacy of the additive model is worth any conceptual incoherence in the genotype-phenotype map (see for instance Box 2 of [225]).

There is one instance in which the importance of additivity is unquestionable. This is evolutionary genetics. Fisher's Fundamental Theorem remains our best model for how the response to selection actually works, meaning that in the absence of genetic

variance that can be described as additive ($h^2 \approx 0$), natural selection will be ineffective [290, 283], as can be seen in Equation 1.4. By this argument, the simple fact of biological evolution argues that additive genetic variance is substantial. If we believe the logic above, the nominally additive variance term is contaminated with effects that are causally epistatic, but it nonetheless must behave roughly according to additive expectations to be selected. Moreover, in cases like those suggested by Wright [293, 294], the fixation of some interacting alleles will give rise to apparent additivity of their interactors in a non-infinite population. Under these circumstances, the idea of additive genetic variation in evolution has little to do with additivity at the physiological level.

1.2.2 Statistical, physiological, and molecular epistases.

As suggested above, even in the specific case of epistasis signifying gene interactions controlling organismal traits, the word can have multiple meanings [207]. 'Molecular epistasis' is the simple case where it has been shown, for instance through synthetic combination of single gene knockout mutants [47], that there is an interaction between genes. However, this synthetic genotype and these alleles may never have existed in nature, and thus they signify a mechanistic link rather than a source of phenotypic variation in natural populations. 'Statistical' and 'physiological' epistasis can, however, concern natural variation; I describe them superficially below but will otherwise direct the reader to Cheverud and Routland [33] for a lucid description.

Physiological epistasis may be measured by enumerating all genotypes of some set of loci and alleles, constructing these genotypes, and measuring the phenotypes of each genotype. One estimates epistasis from these phenotypes by computing the difference between the phenotype of each genotype and the expected additive phenotype based on an average of other genotypes which share portions of the same genotype [15, 287, 211]. Statistical epistasis may be measured more easily, by simply collecting a population sample of the organism in question, measuring their genotypes and phenotypes, and using a model like Equation 1.3 to estimate the interaction terms by least-squares or

a similar method. Naturally, it is impractical to measure physiological or molecular epistasis for very many loci or alleles, but their estimates are necessarily more accurate [33, 211].

Reference to the thought-experiments above regarding the allele *FBN1** may clarify the differences between these types of epistasis. In case (1), we are capable of detecting physiological epistasis but not statistical epistasis, and in case (2) we are capable of detecting both. Specifically, the *FBN1** allele is presumably private to the modern human population, and not in (for instance) Neanderthals. A statistical association of human and Neanderthal height and genotypes therefore cannot detect epistasis with the Neanderthal genetic background, because population structure has induced linkage disequilibrium between *FBN1** and human alleles such that the Neanderthal genetic background never encounters *FBN1**. However, we could in principle introduce the *FBN1** allele into the Neanderthal genetic background (and the Neanderthal *FBN1* allele into the modern human background) to detect physiological epistasis. In case (2), however, the alleles in question are already segregating in the same populations, making both tests possible (in principle).

It is worth noting that some statistical approaches measure not the actual epistatic effects on phenotypes, but rather the linkage disequilibrium between pairs of alleles [45, 226, 13]. These measures, called 'segregation distortion' or 'allelic association' are not properly measures of epistasis, but rather measures of non-independent evolution between alleles. These methods therefore make the assumption that changes in phenotype due to epistatic interactions lead to differential fitness of different genotype combinations [240], and thereby substitute fitness for the phenotype in question.

Because we know that physiological epistasis is the most accurate measure of epistatic contributions to phenotype, a natural question is whether or not it is important in the cases in which it has been measured. A great deal of work in the last decade has been done examining physiological epistasis, usually within single protein molecules [286, 51, 83, 96, 168, 89]. These studies are more or less unanimous in finding that

mutations of large effect are context-dependent, or induce the context-dependence of other mutations.

Moreover, the quantitative analysis of physiological epistasis across these cases estimates very large effects for higher-order interactions (epistasis involving more than two loci) [15, 287, 211]. This finding suggests that the reduction of epistasis even to pairwise interactions misses a large fraction of the mechanistic story. Interestingly, ignoring higher-order epistasis can lead to errors in estimating not only the magnitude but also the sign of lower-order effects in mutational data [15, 211].[4]

1.3 Epistasis in the evolutionary process

The analysis of epistasis' effects on evolution has also been a popular topic. Specifically, many have used experimental evolution as a tool to study how mutations interact in generating phenotypic novelty, and these have generally recognized a large controlling effect of epistasis on the pathways available to evolution [52]. This was largely anticipated by the intuition of Maynard Smith [162], who compared evolution to replacing the letters of a word one by one, where each replacement yields a comprehensible word. Differential fitness in this case is largely determined by 'epistasis', because a misspelling at any position leads to inviability of the whole word. We can think of this case in analogy to a protein 'space', containing all possible proteins of a certain length, and proteins are separated by the number of single substitutions it would take to transform one into another. However, only a subset of proteins (or genotypes, to extend the analogy) are functional (or fit), and thus only a subset of the space can be explored by evolution. This has provided an extremely powerful metaphor for the interpretation of sequential evolutionary data in terms of epistasis [83]. This idea of a restricted phenotypic space has also been studied in purely theoretical systems, which have tended to support Maynard Smith's logic (for instance, [20]).

---

[4] This is more or less the same criticism leveled against ANOVA's method of estimation many decades ago, which was then dismissed as an implausible edge case [302, 283].

The idea that there are relatively few accessible paths in evolution is attributed to the effects of 'historical contingency', meaning the effects of past states on the present state [86]. That is, any given genetic system still encodes the effects of past adaptations, whether or not they have specific function in present circumstances, which in turn affect what adaptations are available in the future, because new adaptations must be attainable to the current system.

1.3.1 Fast-evolving genetic elements

A crucial parameter of the population genetic and evolutionary genetic models discussed above is the mutation rate. Though this rate is generally considered as a single parameter per genome, in reality there is substantial variation within and among genomes for the rates of both nucleotide substitution [140] and more dramatic genomic alterations [123]. Though the importance of fast-evolving genetic elements for adaptation has been discussed substantially in different contexts for many years [177, 122, 77], until recently little attention has been paid to how the segmentation of genomes into high and low mutation regions affects evolutionary outcomes and the genetic architecture of traits. For my doctoral work, I used these classes of variation as models for understanding the relevance of genetic interactions in the two contexts: (1) the genotype-phenotype map, and (2) the evolutionary pathways. We may reasonably expect that mutation classes with large effects per unit time are relevant for understanding both of these topics.

Short tandem repeats

The first part of this thesis, Chapter 2 through Chapter 6, concerns short tandem repeats (STRs), a common and often-ignored class of genomic variation. These elements, which consist of repeated units of 2-10 nucleotides, are best known as neutral markers of genetic variation. This application is facilitated by their high mutation rate from slip-strand error during DNA replication and errors in recombination [77], which can

approach 10^{-2}/generation [277, 77], as opposed to 10^{-9}/generation for nucleotide substitutions. STRs show a distinctive genomic localization pattern, tending to reside in gene promoters [277, 278], in the coding sequences of transcription factors, and in coding regions of genes associated with specific regulatory roles in development [77]. This localization places STRs uniquely as potential contributors to adaptive evolution [122]. Given that many systems emphasize the potential of compensatory evolution (the effect modification of a mutation in the genetic background) to dramatically increase fitness [173, 213, 268], the high mutation rates of STRs suggest that they might preferentially accumulate genetic interactions through compensation [260]. In the long run, this coadaptation by compensation would lead to genetic incompatibilities and potentially speciation [192].

However, genotype-phenotype maps to date are developed largely as functions of biallelic single nucleotide variants, not for multiallelic STRs. Certain work in the past suggested that a sufficient simplifying model of STR variation is to treat it as a linear input to a monotonic function of phenotype [236, 277], suggesting that simple additive linear models can be adapted to describing the effects of STR variation on phenotypes.

In collaboration with others, I use Chapter 2 to study the STR genotype-phenotype map problem in the case of a naturally highly variable coding STR in the *ELF3* gene of *Arabidopsis thaliana*. We used transgenic analysis in two genetic backgrounds of *A. thaliana* to measure the physiological epistasis of these allelic arrays, which indicated strong non-additive effects of *ELF3* STR variation. Again in collaboration, in Chapter 3 I study a case of another coding STR in the *PFT1* gene in *A. thaliana* under what appears to be stabilizing selection, and used the mapping of artificial variation to phenotype to attempt to understand the basis of this selection. In Chapter 4, I discuss these cases and draw inferences from new high-throughput sequencing technologies to identify STRs potentially under selection or associated with phenotypes. I predict that highly variable STRs such as *ELF3* might be best understood in light of potentially epistatic genetic architectures of complex traits. In Chapter 5, I follow this intuition to map the

genetic basis of the epistasis observed for the *ELF3* STR. In Chapter 6, I consider the role of environmental effects interacting with the *ELF3* STR to generate phenotypes, and discover a previously unknown role for ELF3 in temperature sensing.

Horizontal acquisition and deletion of genetic material

The second part of this thesis, comprised of Chapters 7 and 8, concerns another class of non-substitutional mutation with outsize effects on evolution: horizontal gene transfer (HGT). In this process, large pieces of DNA are exchanged between organisms (one of which is typically a prokaryote). While the rate of these events is not much higher than the rate of substitution (though it is much higher than the rate of gene duplication [272], for instance), the effect of each event is probably massively larger [219, 128]. Because the rate of DNA deletion in prokaryotes is also generally very high [167, 136], on average prokaryotes need to acquire quite large amounts of DNA simply to maintain a stable genome size, and are consequently undergoing constant genomic flux [219].

There are many constraints on the kind of DNA that acquired and maintained through evolutionary time, which one might otherwise expect to be lost [10]. First, for homologous recombination to integrate exogenous DNA, it must generally come from a relative with fairly low sequence divergence [158]. Second, there must be sufficient DNA repair machinery to effect such a recombination [214]. Third, certain classes of genes are more likely to transfer than others [111]. Fourth, horizontally acquired gene products may interfere with proteostasis [17]. Fifth and last, there is some evidence that functional interactions with the host genome (e.g. metabolic links or protein cofactors) may bias the kinds of genes that are acquired [199, 109]. All of these constitute mechanisms by which interactions with the host genome can constrain the future evolution of genomes. A separate question is whether these lead to global patterns in genome evolution, or whether these constraints are insignificant relative to the role of chance and direct selection on specific single genes (for example, for antibiotic resistance [243]).

In Chapter 7, I start with the expectation that genes do interact in their evolution, specifically focusing on the case of bacterial Hsp90, using some methods of measuring coevolution or coordinated evolution mentioned above. In eukaryotes, Hsp90 is well-known as a hub of genetic interactions [269, 137], due to its role as a promiscuous molecular chaperone of other proteins. Specifically, I search for prokaryotic genes showing coordinated gains and losses with Hsp90, to explore whether its functions in Prokarya reflect the known diversity of Hsp90 dependencies among Eukarya. In Chapter 8, I describe a comparative genomics approach designed to directly evaluate whether the genome-constrained acquisition of new genes in Prokarya leads to predictable evolutionary patterns. Again assuming that covariation of genes indicates functional links, I used the directionality of evolutionary change noted by Maynard Smith to infer evolution-controlling dependencies.

Using this combination of approaches, both experimental and analytical, I attempt to demonstrate how we can use our knowledge and our expectations about the distribution of epistasis to inform genetic and evolutionary research.